

[19]中华人民共和国国家知识产权局

[51]Int. Cl<sup>7</sup>

H04N 7/15  
H04N 7/14

[12] 发明专利申请公开说明书

[21] 申请号 98805462.0

[43]公开日 2000 年 6 月 21 日

[11]公开号 CN 1257631A

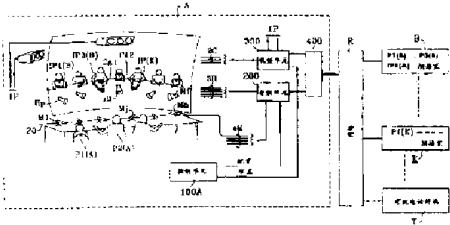
[22]申请日 1998.3.26 [21]申请号 98805462.0  
[30]优先权  
[32]1997.3.27 [33]FR [31]97/03785  
[86]国际申请 PCT/FR98/00625 1998.3.26  
[87]国际公布 WO98/44734 法 1998.10.8  
[85]进入国家阶段日期 1999.11.25  
[71]申请人 法国电讯公司  
地址 法国巴黎  
[72]发明人 G·布赫纳 Y·马约克斯  
C·维普利兹  
J·-P·朱利安

[74]专利代理机构 中国专利代理(香港)有限公司  
代理人 王 勇 张志醒

权利要求书 3 页 说明书 9 页 附图页数 9 页

[54]发明名称 视频会议系统  
[57]摘要

本发明涉及位于不同场地与会者之间的视频会议系统,每一场地 装备一个观看屏幕。根据本发明,观看屏幕为投影类型、高射投影类 型或直接投射类型的大屏幕 (10),包括一个或多个形成一个图像壁的 观看系统,以使同时在由所述屏幕(10)形成的图像壁上产生的一个 空间内观看所有远处的与会者。所建议的系统包括为每一场地在空间 分布声音的设备(100A,100K,Hi,Mj),以 匹配由每一远处的与会 者发出的声音和在所考虑的场 地的屏幕上投射的该与会者的图像。



ISSN 1000-84274

## 权 利 要 求 书

1. 一个在位于几个不同场地的与会者之间的视频会议系统, 其包括在每一场地的一个观看屏幕、多个摄像机和图像再现设备以及声音  
5 记录和再现设备, 其特征在于:

- 观看屏幕是投影、高射投影或直接投影类型的大屏幕(10), 包括一个或多个形成一个图像壁的观看系统, 以便在由所述屏幕(10)形成的图像壁上产生的一个空间内同时观看所有远处的与会者,

- 声音记录和再现设备, 包括在每一场地空间分布声音的设备  
10 (100A, 100K, Hi 和 Mj), 以匹配由每一位远处与会者发出的声音和在所考虑场地投射的该与会者的图像,

- 一给定场地接收由其它场地发布的图像和音频信号并根据建立  
15 的对应关系再现它们, 通过在两地之间进行话音交换实现同时观看和聆听其它场地和发言者, 无缝地且并不干扰其它与会者.

2. 根据权利要求1的视频会议系统, 其特征在于, 形成图像壁的  
20 屏幕(10)的大小如此选择, 使得能够显示所有场地的图像, 而所显示的与会者的比例接近1.

3. 根据权利要求1的视频会议系统, 其特征在于, 拍摄设备包括  
25 一个初始摄像机(CA1)和一个第二摄像机(CA2), 其中记录的声音与  
所述摄像机的一个或另一个关联, 这取决于所传输的视频信号来自所述摄像机(CA1, CA2)中的哪一个.

4. 根据权利要求1的视频会议系统, 其中, 声音记录设备包括几个  
30 麦克风、音频信号捕获设备(201)以及为所述信号加密和阵列化的设备(205).

5. 根据权利要求1的视频会议系统, 其特征在于, 声音再现设备  
25 包括对音频信号解码和解阵列的设备(203), 再现音频信号的设备  
(205)和几个扬声器.

6. 根据前述权利要求中任何一个权利要求的视频会议系统, 其特征  
30 在于, 声音记录和再现设备包括:

- 信号和数字化捕获单元(201);

- 信号再现和模数变换单元(205);

- 在所述场地的与会者面前放置的n个麦克风(M1... Mn);

- 沿屏幕长度放置的  $p$  个扬声器 ( $H1 \dots H_p$ ), 其中  $p$  正比于屏幕大小;

- 包括位速率压缩设备的网络适配设备;

- 回声控制设备。

5        7. 根据前述权利要求中任何一个权利要求的视频会议系统, 其特征在于, 提供在麦克风、由所述麦克风提供的信号和打算再现所述信号的扬声器之间匹配的设备包括存储建立所希望的匹配的一个或多个配置的存储设备 (110)。

10       8. 根据前述权利要求中任何一个权利要求的视频会议系统, 其特征在于, 对于每一场地每一拍摄单元包括  $q$  个摄像机, 它们放置在所述场地与会者面前, 使得每一个摄像机拍摄显示在屏幕不同部分的不同组的外地与会者。

15       9. 根据前述权利要求中任何一个权利要求的视频会议系统, 其中在  $N$  个场地之间召开视频会议, 其特征在于, 每一场地的摄像机数  $q$  大于或等于 2, 以再现图像透视效果。

20       10. 根据权利要求 1 到 8 中任何一个权利要求的视频会议系统, 其中只在两个不同场地之间召开视频会议, 其中, 一个摄像机 ( $CA1$ ) 拍摄包括一初始与会者组的本地场地的部分图像, 另一摄像机 ( $CA2$ ) 拍摄包括第二组与会者的本地场地的部分图像, 这两部分图像并列放置, 使能看到所有与会者。

11. 根据权利要求 1 到 10 中任何一个权利要求的视频会议系统, 其中使用几个摄像机拍摄, 其特征在于, 所述的摄像机 ( $CA1, CA2$ ) 放置在投射的图像下面。

25       12. 根据前述权利要求中任何一个权利要求的视频会议系统, 其特征在于, 所述系统还包括若干阴极射线设备、或液晶视频、或高射投影仪 ( $PR1, PR2$ ), 用于以接近 1 的比例再现图像。

30       13. 根据前述权利要求中任何一个权利要求的视频会议系统, 其特征在于, 所述系统包括用于每一场地的图像处理设备, 能够并置或叠加从外场地接收的图像 ( $301, 304$ ), 以使所有图像能在屏幕 (10) 上显示。

14. 根据前述权利要求中任何一个权利要求的视频会议系统, 其中, 场地经由不同的高位速率传输网络 ( $R$ ) 直接连接, 或经由一个视



视频会议网关或经由一个网络和视频会议网关间接连接。

15. 根据前述权利要求中任何一个权利要求的视频会议系统，其特征在于，处理设备包括：

- PC 类型或工作站（100A...，100K）的处理和控制单元，

5       - 由处理单元控制的视频编码器 - 解码器（303 - 306）和音频编码器 - 解码器，

- 用于所述网络的传输适配接口（301，302，204）。

16. 根据权利要求 15 的视频会议系统，其特征在于，所述网络是一个 ATM（异步传输模式）类型的传输网络，用于音像信号的传输接口是 AAL1 类型的接口，在处理单元之间用于视频会议的控制信号的传输接口是 AAL5 类型接口。

10

# 说明书

## 视频会议系统

5 本发明涉及在位于远处、通常在外地的与会者之间的一个视频会议系统或者电信会议系统。

它适用于多点系统，亦即为产生在位于几个地点的与会者之间的视听链路而设计的系统。它还涉及点对点系统，亦即通过一条视听链路只连接两个地点的系统。它还适用于经由一个会议网关连接的可视电话 (visiophones)。

10 用于一给定地点的视频会议系统的设备安放在一个设施中，以下称为房间或演播室。在大多数场合，与会者围坐在一张会议桌前，面对一个显示屏。

今天，商用视频会议服务和系统提议一个有限音像质量的多点链路，这或者是由于质量信道限制 (banks) (声音和图像恢复)，或者  
15 本质上是由于所用特定网络通带的限制而没有位速率的事实所致。

的确，会议系统根据处于点对点方式还是多点方式等几种配置连接到数字网络，主要是 ISDN 网络。

所以，在 NUMERIS 网络上为此种服务提供的位速率在 128 千位 / 秒和 384 千位 / 秒之间变化，前者用于下限范围会议服务，后者用于  
20 上限范围会议服务。

所用终端大体符合 ITU 标准，诸如所有 H320 标准。

在某些系统中，可能在某一时刻只看见远处的一个房间。这使用户很不方便，因为他们不能同时看见每一个人。手动或自动切换选择投射到屏幕上的房间。一般来说，这一房间发送最强的音频信号 (声音切换)。  
25 会议网关也是这样，它根据声音的检测、亦即具有最活跃声音的房间切换图像。

遵守 ITU 标准、诸如标准 320 或 H323 或者其它标准的其它系统使用一个中心单元从所有多会议单元 (MCU) 接收图像，正如在标准 H231 和 H243 中规定的那样。这一单元通过共享给 MCU 传输的通带恢复图像，其次数和所发送的图像数目同样多。  
30

这一点通过加密实现，其基本上以一个在 40 到 50 之间的压缩因子压缩连机位速率。这导致清晰度损失等于大约由每一房间 (多点系

统多达 5 个房间) 所传输图像的四分之三。

图像的数字压缩可能为运动型 JPEG(ISO 标准)或 MPEG1 或 MPEG2 的以 H261 加密的 H320 类型, 或者为以 H263 加密的 H323 类型。

在上述任一种中, 发送的图像都可以采取计算机文件的形式。

- 5 在所有情况下, 都会削弱接收到的图像, 不可能重现多点会议的效果。

对于音频传输, 使用加密系统或者 G722 类型或 G711 类型的位速率压缩, 它不考虑话音信号的原始质量, 无论是就带宽还是就加密质量自身。

- 10 根据本发明的视频会议系统的目标是在若干远程场地之间提供最大数目与会者的视频会议, 超过当今商用系统的限制。它给位于远处的与会者在屏幕上提供视图(以接近于 1 的比例), 使清楚感知行为的特征和方式, 和使声音和图像匹配, 这是由于根据成像条件在空间分布远处与会者的声音而实现的。

- 15 建议的系统甚至在多点配置下也连续地显示在远程场地的与会者。此外, 该系统通过使用 N 个接近要拍摄的图像放置的(例如安放在屏幕窗口内的)视频摄像机使眼睛的接触损失为最小, 后面将会说明。它还能再现高保真声音。

- 20 更具体说, 本发明建议一个在若干远程场地的与会者之间的视频会议系统, 每一场地安置一个视频屏幕, 其中, 视频屏幕是视频投影仪类型、高射投影仪(overhead projector)类型或直接视频类型的大屏幕, 包括一个或者多个形成图像壁的视频系统, 以便在由所述屏幕形成的图像壁上再现的一个位置处同时观看所有远处的与会者。

- 25 - 所述系统为每一给定场地包括在空间分布声音的设备, 以匹配由每位远程与会者发出的声音和在所考虑的场地的屏幕上看到的他/她的图像,

- 30 - 从而一给定场地接收由其它场地发布的图像和音频信号, 并根据建立的对应(correspondence)再现它们, 观看与聆听其它场地和发言者是同时的, 并基于同时在几个场地之间进行的话音交换, 无缝地且不干扰其它与会者。

远处与会者的图像以接近于 1 的比例显示在演播室屏幕上; 该比

例取决于屏幕和与会者围坐的桌子之间的距离。例如，这导致一个近似为 5X2m 的标准大小的屏幕来观看 4 个在每一房间约有 4 人的远处的房间。

5 该视频会议系统包括一个或多个摄像机 (CA1)、(CA2) 和声音记录系统 (麦克风或声学天线)。声音记录数据与一个或另一个摄像机匹配，这取决于所传输的视频信号来自摄像机 (CA1) 还是摄像机 (CA2)。

更具体说，这一匹配提供一种“主观叠加” (subjective overlay)，声音源接近关联的图像 (沿图像的轴线)。

根据本发明另一特征，声音记录和再现设备包括：

- 10 - 信号捕获和数字化单元，
- 信号再现和模数变换单元；
- 在所述场地的与会者面前分布的 n 个麦克风；
- 沿屏幕长度分布的 p 个扬声器，其中 p 正比于屏幕大小；
- 在一个或多个麦克风、从所述麦克风发出的信号和用于再现所述信号的外地扬声器之间的匹配单元，
- 15 - 装备有位速率衰减的网络适配设备；
- 回声控制设备。

20 声音在空间分布，以便使声音与视觉图像匹配。这种布局不仅强烈增强电信会议效果，同时能使在两个远处的房间中并行进行若干对话；对话的监视通过系统能够注意于某人希望听的人员而简化，就像在正常的会议中一样。

在麦克风、来自所述麦克风的信号和用于再现所述信号的远程场地的扬声器之间建立对话的设备通过编程希望的配置而操作。这种编程可能涉及存储一个或者多个预先确定的配置。

25 根据另一特征，图像记录器包括为每一场地的 q 个摄像机，这一数目优选超过或等于 2。这些摄像机放置在所述场地与会者面前，使其位于在屏幕上显示各个外地与会者的不同区域或接近这些区域。在实际中，这些摄像机可以位于在屏幕的下面、顶部或周围。

30 因此，每一房间可以包括几个摄像机，它们以不同角度拍摄与会者。传输到某场地的图像是由位于接近为该场地投射的图像的摄像机记录的图像。这使得可能减小眼睛接触效果，区别视点和根据他/她在该场地和在总配置中的位置重新建立每一与会者的位置。

在几个图像沿屏幕宽度并置的点对点配置中，每一图像提供该设施的部分视图。这使得容易知道在屏幕上谁在看谁，从而跟踪会议动态。

在实际中，摄像机放置在屏幕 10 上投影的图像的下面。

5 在具有两个相连图像（一个在另一个的旁边）的配置中，摄像机重新排列和放置在距屏幕中心一定距离处，相应于该图像宽度的约三分之一，亦即屏幕总宽度的六分之一。这样的重排列使由两个摄像机拍摄的两个图像在边缘处的重叠的问题为最小。

10 多个场地在点对点或多点配置中通过一个高位速率网络或者通过一个中央 MCU 单元（视频会议网关）连接在一起。

本发明的其它特征和优点在阅读这里的说明书时会显现。本说明书参考附图以非穷举的解释给出，附图中：

图 1 和图 2 原理说明本发明解决的远程会议问题，

图 3 表示本发明的原理图，

15 图 4 表示在一个演播室中的设备的实施例的图，

图 5 是相应于在 3 个演播室之间有一条会议链路的场合下的一个特定实施例的图，

图 6 是相应于在两个演播室之间有一条链路的场合下的一个特定实施例的图，

20 图 7 到图 9 表示形成大屏幕上图像壁的不同“屏幕图像”的一个例子，

图 10 到 13 表示在演播室设备或视频话音终端之间根据本发明在不同网络配置中的不同连接方式，

25 为说明实现远程会议效果的一种方法，图 1 表示出类似真实生活中的会议形象。

假定有 12 个人围坐在一张虚拟的圆桌前。一个与会者可以通过一个窗口看到该组另外 4 个相邻的人，并且通过开发表面，可以假定该与会者可以在一个平面或几乎纯平的屏幕上看见这 4 个人。因此他将在两个特别是以平面结构的屏幕上看见两个不同演播室中以 4 人为一组的其它 8 个人。作为一般规则，虚拟桌不是圆形，而是纵长放置的，与会者小组分坐在两边。因此，人 E 将看见 A 在他/她的左边，而 F 在他/她的右边，所有出席的个人连续相邻排列。

30





或与其表面)。

对于围坐在桌子周围的与会者，在给定场地的会议桌前安置一个大屏幕 10，(每一场地的) 屏幕图像都投射在该屏幕上，使得远处的与会者和在该场地的与会者处于同一水平，从而产生一张虚拟桌。

5 当与会者 P1 (A) 对与会者 P1 (B) 说话时，场地 A 的摄像机 CA1 的图像被发射到其它场地 B 到 K。此外，场地 B 的摄像机 CA1 的图像被发送到场地 A、C 到 K。如果在其它场地无人发言，则摄制这些场地的摄像机不反应，尽管还是发送所有与会者的图像。

10 摄像机的图像由位于各场地处的控制单元 100A - 100K 控制，或由一个监控场地遥控。

在声音记录方式和必须记录图像的摄像机之间的伺服控制由熟悉本技术领域的人所公知的控制程序保证。该程序由单元 100A 到 100K 实现，以便获得这样的图像，它们一旦被投射在屏幕上便产生所涉及的人确实正在看另一个人的印象。

15 在实际中，这一效果可以只用两个摄像机产生。

由网络传输的图像被每一场地接收，并由视频单元 300 处理，以便被并列和 / 或重迭投射到屏幕上。然而，所有接收的图像被投射到每一场地的屏幕 10 上，所述屏幕形成一个图像壁。

20 在点对点系统的特例下，投射到每一屏幕上的图像相应于单一房间。这一配置有可能加强图像质量。为此目的，建议传输两个图像，各分别由一个摄像机拍摄。每一摄像机记录该设施的部分图像。实际上，一个摄像机拍摄场面的一半，而另一摄像机拍摄另一半。

25 图像的并置由处理单元 100A 和 100B 控制，它们根据由这两个摄像机发送的信号控制视频处理单元在屏幕上投射单幅图像。图像的级连在本质上是传统类型的图像处理。

图 4 更完整地表示作为系统部件的处理和控制单元。这又是在每一演播室有两台摄像机的另一个例子。

对于视频部分：

30 第一摄像机 CA1 连接到视频编码器 303，而视频解码单元 304 连接到初始投影仪 PR1。解码单元包括一个或者多个解码器，使得每一个解码器解码来自该场地的视频信号。在下面的说明中，将讲述解码器 304，因为可适用于一组解码器。

编码器和解码器连接到网络 R 的视频信号适配和传输接口 301。  
例如，所述接口是用于 ATM 网络传输的 AAL1 接口（适配层）。

上述接口连接到一个物理支持接口 400。

例如，控制单元 100A 根据该视频会议配置经由本地网络 RL 分别为编码器、解码器和适配接口 301 发送控制信号 CD1、DC1、IV1。

相似地，第二摄像机 CA2 连接到编码器 305，编码器 305 自身连接到网络的一个传输适配接口 302。解码器 306 连接到第二投影仪 PR2 和接口 302。接口 302 连接到物理支持接口 400。

控制单元根据该配置分别为编码器、解码器和适配接口发送控制信号 CD2、DC2、IV2。

可以把一个视频编码器在功能上连接到几个解码器，此种技术不属于本发明的范围。

然后可以把一个场地的第三摄像机连接到编码器 303 或 305 之一，并发送优先考虑的某特定人的图像。这一摄像机的图像将由该摄像机连接的编码器加密，并传输到所有其它场地。然后所有其它场地必须在解码器单元 303 或 306 中装备一个附加的解码器，以处理由该摄像机提交的视频信号，并投射该拍摄的图象到与这种图像关联的场地的屏幕 10 上。

投影仪可以是视频投影仪，也可以是液晶高射投影仪（liquid crystal overhead projector）。只要使用等离子屏幕的话，可以直接投射。

对于音频部分：

n 个麦克风连接到一个信号捕获单元 201，它把信号数字化。该单元 201 的输出加在衰减单元 207 的输入以控制回声。衰减单元的输出加在位速率缩减器编码器的输入用于由单元 100A 控制的音频信号 202。编码器 202 还执行麦克风信号的阵列化，以与网络上可用的用于音频传输的通道数目一致。为此目的，编码器接收配置控制信号 CA。音频适配编码器 202 连接到网络 R 的传输适配接口 204，其是 ATM 网络的 AAL5 接口。该接口连接到物理支持接口 400。

接口 204 能在网络上发布加密的音频信号和接收由其它场地发布的加密的音频信号。接收的信号由接口 204 处理并加在一个解码单元和音频信号解阵列单元 203 上。

由解码和解阵列单元 203 提交的信号经由一个音频信号再现和模数变换单元 205 施加在希望的一个（或多个）扬声器上。该单元 205 设计用于发送接收到的需要的信号到几个适当的扬声器。

在配置被编程且被存储在控制单元的存储器 110 中的场合，单元 100A 将在其存储器表中包含有关信息，指示解码器的输出 S1 正在提交信号，该信号由麦克风 M1 或包括场地 B 的 M1 的一系列麦克风拾取，以及该输出随后必须被传送到扬声器 Hn 或包括 H1 的一系列扬声器，以产生发言的与会者的一个特写镜头。

选择捕获系统（麦克风和单元 205）要注意话音信号通带，并以 7KHz 的频带（对质量而言是该系统的下限）操作。同样选择加密系统以注意这一原始质量，同时选择位速率缩减算法，如 MPEG1 或 MPEG2 标准所说明的。还选择捕获系统以提供一个好的信噪比。

在点对点配置的两个房间的场合，一个可能的配置包括 6 个声音通道，亦即为每一个麦克风一个通道。在 3 个房间的场合，在每一场地有 2X3 个声音通道。对于场地 A，3 个通道提供 A 和 B 之间的音频链路，而另外 3 个提供 A 和 C 之间（类似地为 B 和 C 之间）的音频链路。

以最低可能的时间位移（约 100 毫秒）传输图像和声音，结合视频图像的质量和大小，它将允许聋人对外地发言人进行唇读。

此外，音频设备包括一个回声控制电路 206，其通过本质上是标准处理的可变衰减单元 207 用作对麦克风信号的一个增益变换器。所述电路连接到单元 201 的输出端和单元 205 的输入端。

图 5 表示上述的一个可能配置。这是在 3 个场地 A、B 和 C 之间的视频配置。

所述图通过强调该配置在不同场地之间提供的直接链路而忽略传输网络而加以简化。

从 3 个场地 A、B 和 C 来说明这一配置。

每一场地 A 和 B 各有两个摄像机 CA1 和 CA2，它们以不同角度在同一场地拍摄。

根据这一配置，为编码器和解码器提供点对点链路。

摄像机 CA1 连接到场地 A 的编码器 303。该编码器连接到解码器 304，后者连接到提供屏幕图像 E1（A）（来自场地 A 的摄像机 CA1 的图像）的投影仪。



说明书附图

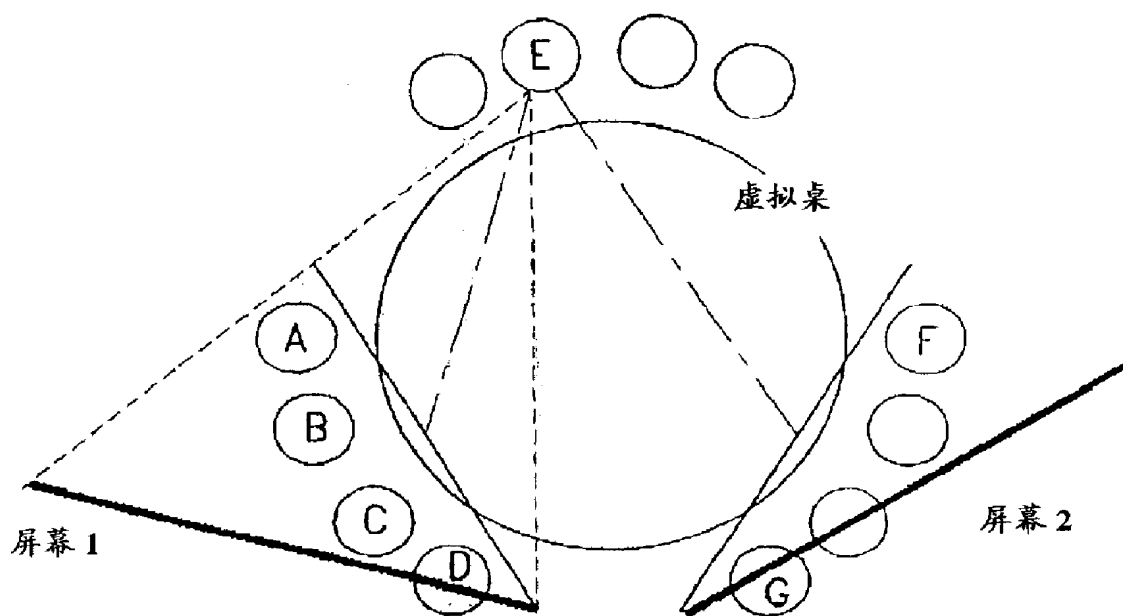


图 1

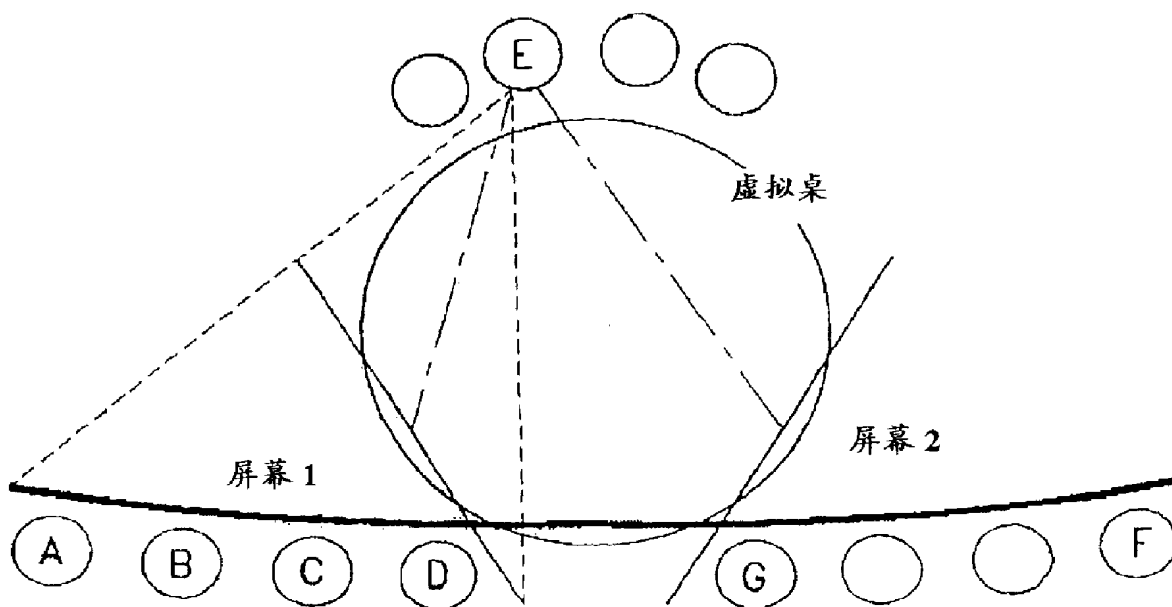


图 2

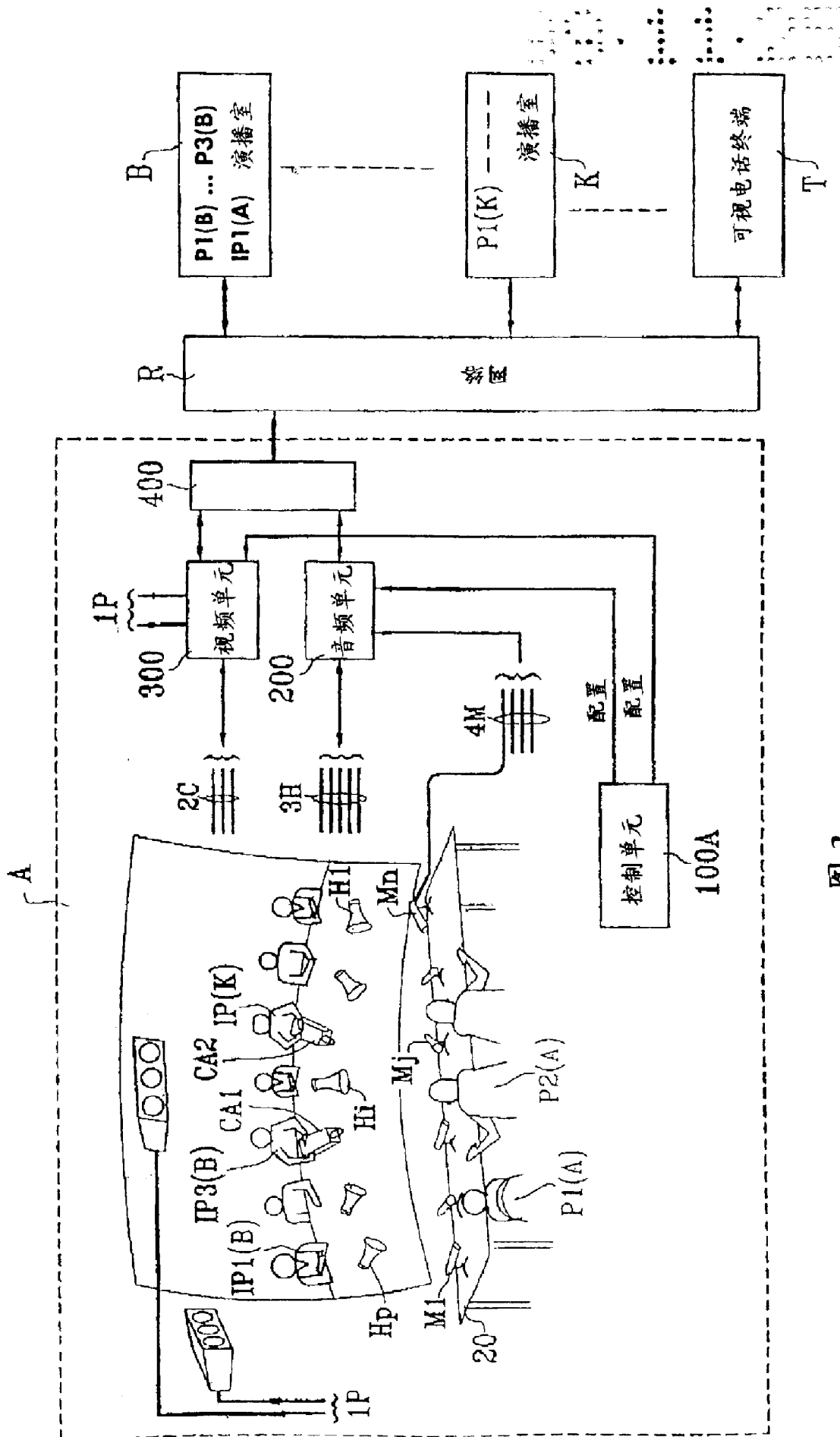


图 3

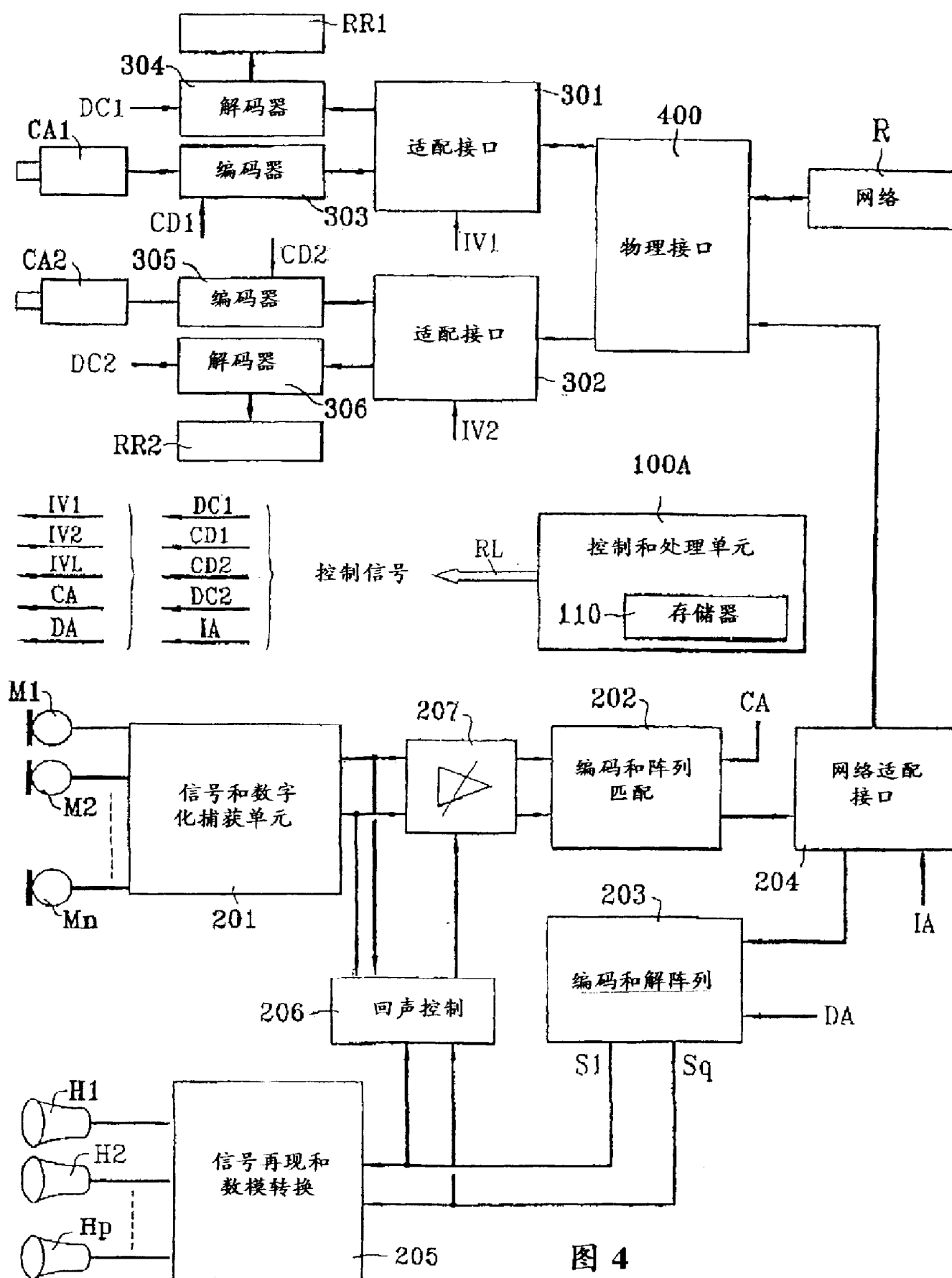


图 4



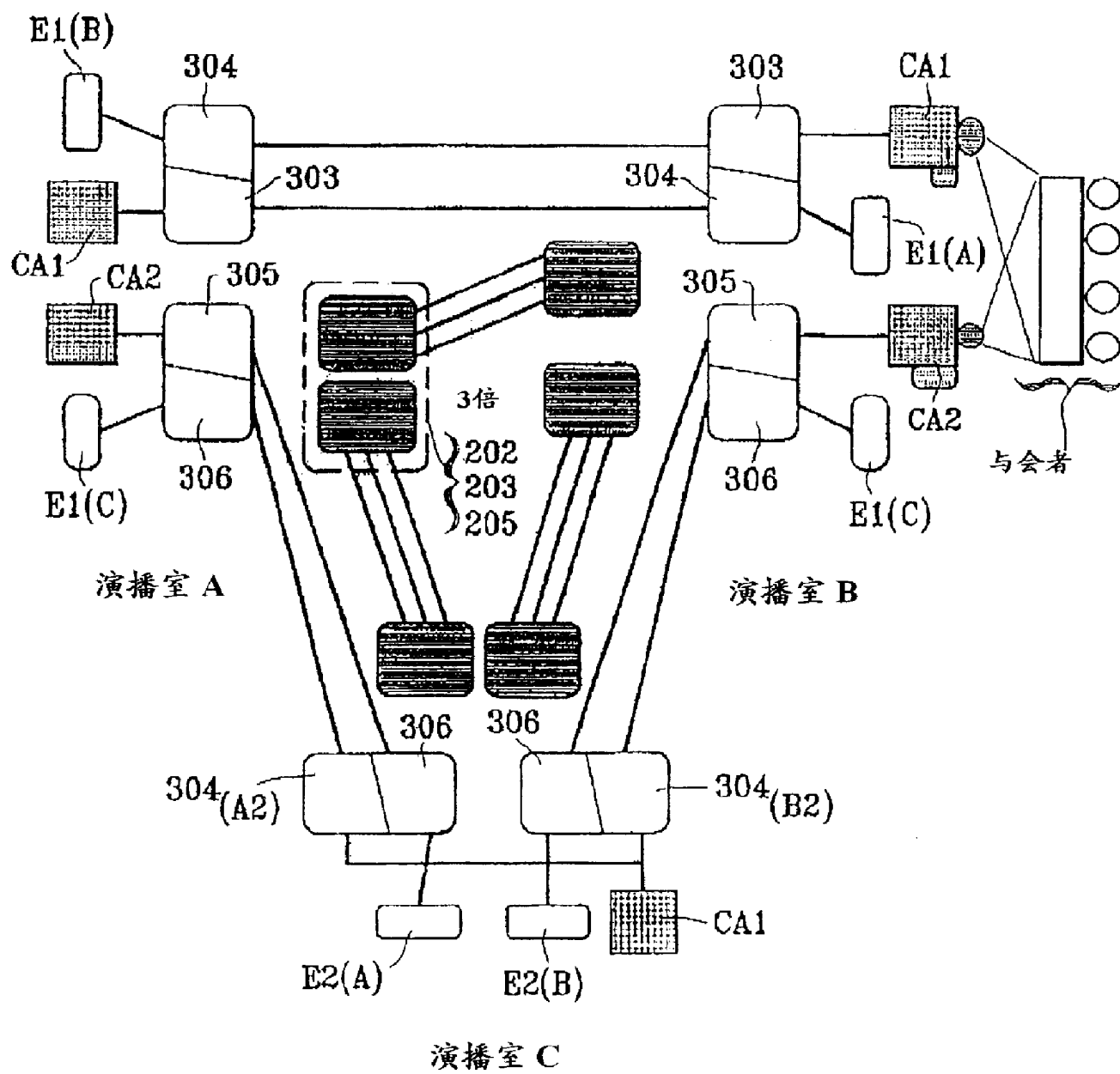


图 5



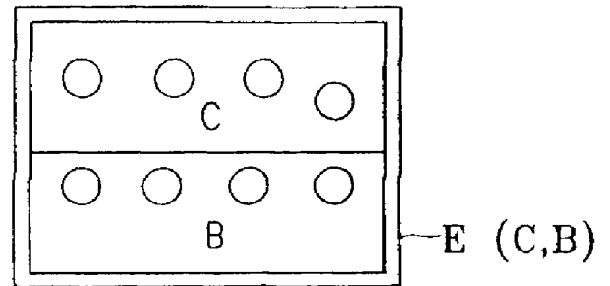
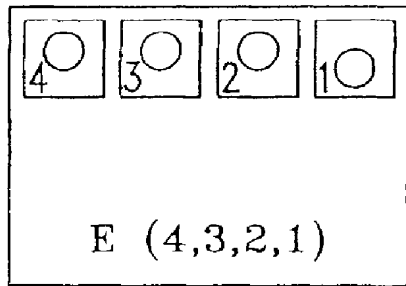


图 7

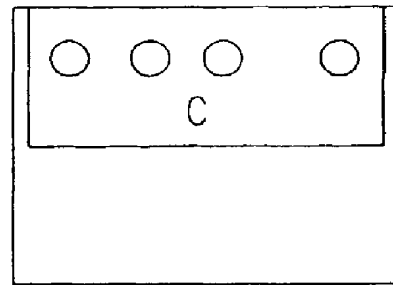
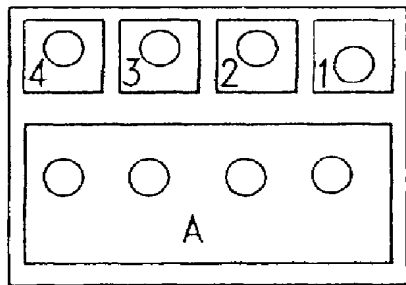


图 8

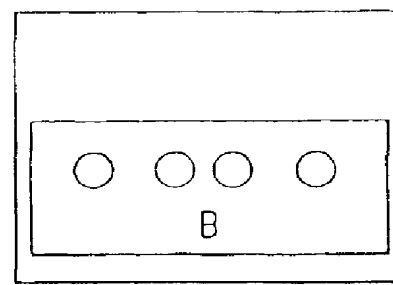
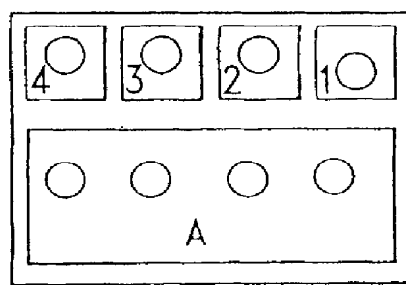


图 9

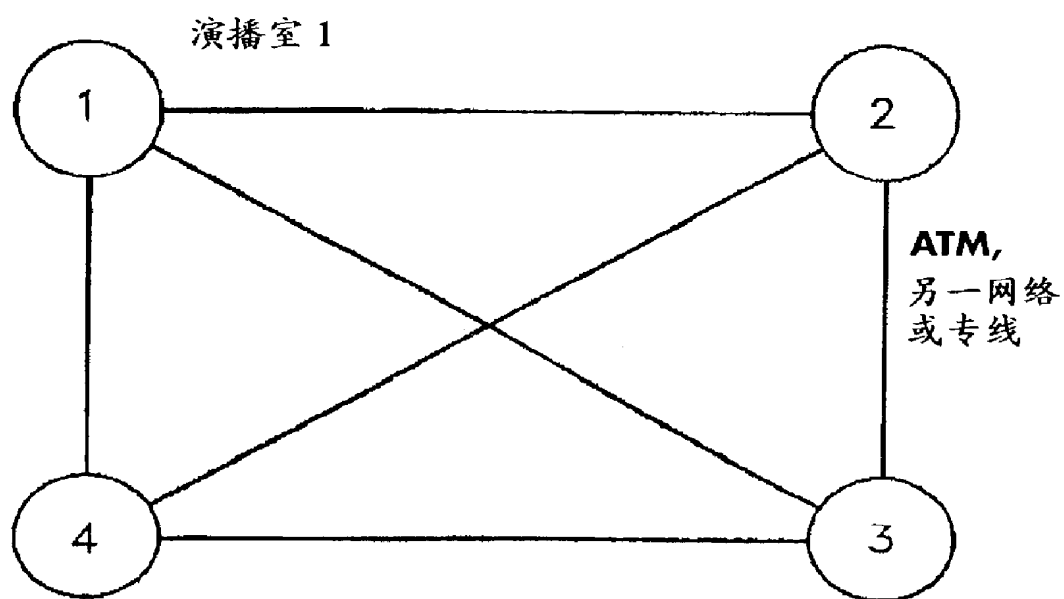


图 10  
网状连线

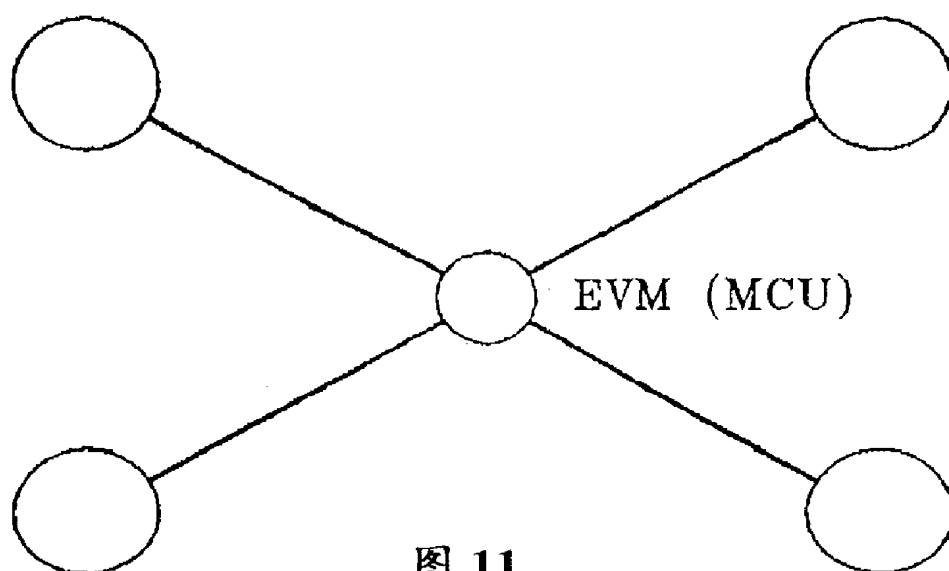
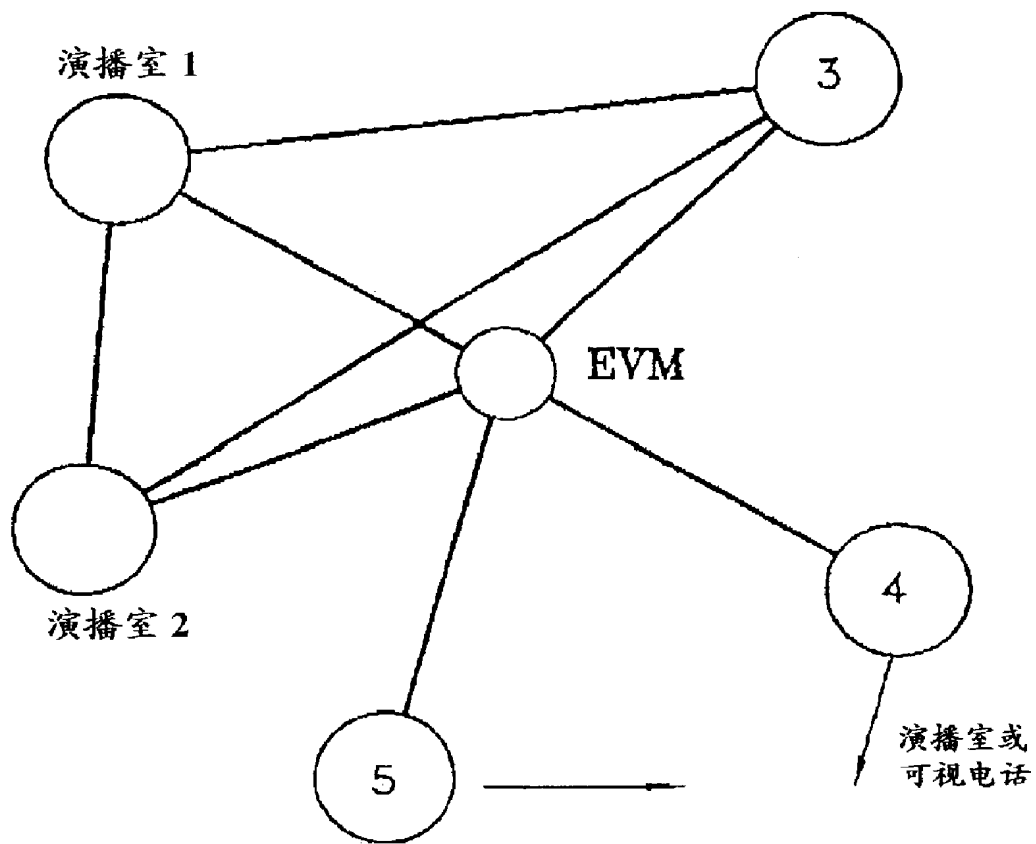


图 11  
经由MCU设备的中心连线



**图 12**  
混合连线配置

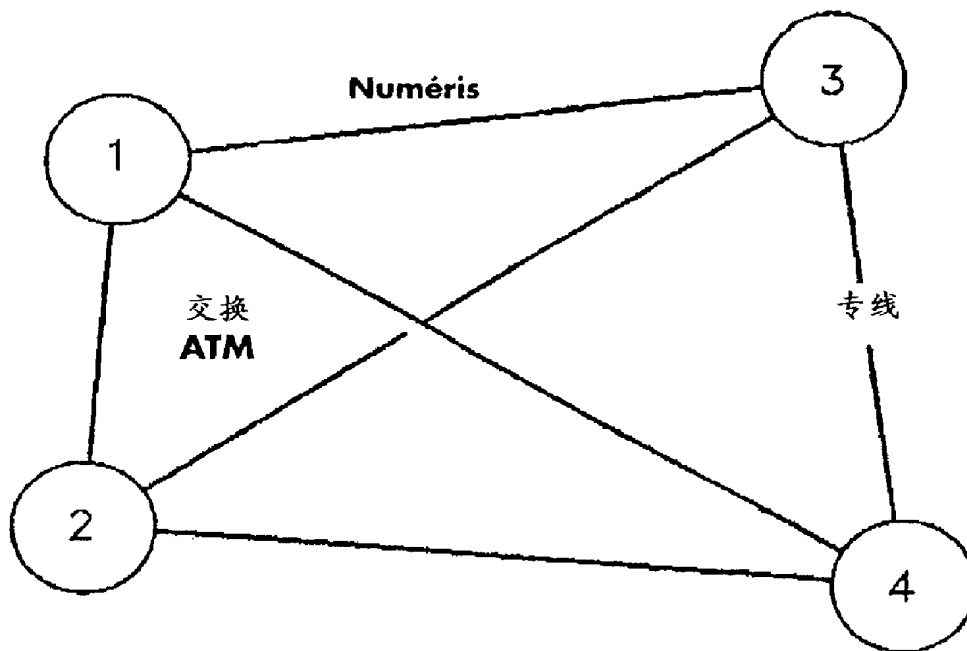


图 13  
不同网络连线